Statistical efficiency of methods for computing free energy of hydration

Yildirim, Ahmet; Wassenaar, Tsjerk A.; van der Spoel, David

Published in:
Journal of Chemical Physics

DOI:
10.1063/1.5041835

Publication date:
2018

Statistical efficiency of methods for computing free energy of hydration

Cite as: J. Chem. Phys. **149**, 144111 (2018); https://doi.org/10.1063/1.5041835
Submitted: 27 May 2018 . Accepted: 27 September 2018 . Published Online: 11 October 2018

Ahmet Yildirim, Tsjerk A. Wassenaar, and David van der Spoel

Statistical efficiency of methods for computing free energy of hydration

Ahmet Yildirim,[1] Tsjerk A. Wassenaar,[2] and David van der Spoel[3,a)]
[1]*Department of Physics, Siirt University, Siirt 56100, Turkey*
[2]*Groningen Biomolecular Sciences and Biotechnology Institute and Zernike Institute for Advanced Materials, University of Groningen, 9747 AG Groningen, The Netherlands*
[3]*Uppsala Centre for Computational Chemistry, Science for Life Laboratory, Department of Cell and Molecular Biology, Uppsala University, Husargatan 3, Box 596, SE-75124 Uppsala, Sweden*

(Received 27 May 2018; accepted 27 September 2018; published online 11 October 2018)

The hydration free energy (HFE) is a critical property for predicting and understanding chemical and biological processes in aqueous solution. There are a number of computational methods to derive HFE, generally classified into the equilibrium or non-equilibrium methods, based on the type of calculations used. In the present study, we compute the hydration free energies of 34 small, neutral, organic molecules with experimental HFE between +2 and −16 kcal/mol. The one-sided non-equilibrium methods Jarzynski Forward (JF) and Backward (JB), the two-sided non-equilibrium methods Jarzynski mean based on the average of JF and JB, Crooks Gaussian Intersection (CGI), and the Bennett Acceptance Ratio (BAR) are compared to the estimates from the two-sided equilibrium method Multistate Bennett Acceptance Ratio (MBAR), which is considered as the reference method for HFE calculations, and experimental data from the literature. Our results show that the estimated hydration free energies from all the methods are consistent with MBAR results, and all methods provide a mean absolute error of ∼0.8 kcal/mol and root mean square error of ∼1 kcal for the 34 organic molecules studied. In addition, the results show that one-sided methods JF and JB result in systematic deviations that cannot be corrected entirely. The statistical efficiency ε of the different methods can be expressed as the one over the simulation time times the average variance in the HFE. From such an analysis, we conclude that ε(MBAR) > ε(BAR) ≈ ε(CGI) > ε(JX), where JX is any of the Jarzynski methods. In other words, the non-equilibrium methods tested here for the prediction of HFE have lower computational efficiency than the MBAR method. https://doi.org/10.1063/1.5041835

I. INTRODUCTION

Hydration or aqueous solvation of molecules is essential in many biochemical processes such as transfer of compounds through the cell membrane or the activity of the biological macromolecules in cells[1] but also in chemical processes, such as micelle formation, protein folding, and aggregation or binding of drugs to biological macromolecules. The hydration free energy (HFE) is the amount of free energy needed to transfer a molecule from the gas phase to aqueous solution. It aids in understanding the outcomes of various chemical and biological processes in aqueous solutions.[2,3] Computational approaches to predict HFE are important to understand molecular interactions in the aqueous phase.

There has been extensive research on the computation of HFE. For instance, hydration free energies have been calculated for 504 small neutral organic molecules in an implicit solvent[4] and explicit solvent[5] using the Bennett acceptance ratio (BAR)[6] based on molecular dynamics (MD) simulations. In a recent study, Matos *et al.*[7] have used the multistate Bennett acceptance ratio (MBAR) to get highly accurate HFEs of organic molecules, and the results were presented in the FreeSolv database of hydration free energies.[8] Furthermore, HFE calculations have been reported for amino acid side chain analogs[9–13] and 60 small molecules.[14,15]

The simulation methods giving accurate HFE results use computationally expensive explicit solvent models and usually take days or weeks to complete a single calculation.[16–18] Faster simulation methods using implicit solvent, such as the generalized Born (GB)[19] and Poisson-Boltzmann (PB)[20,21] methods, are less accurate.[22–25] Gibbs free energies of solvation in other solvents than water are effectively uncorrelated to experimental data with correlation coefficients R from 0.25 (GB) to 0.5 (PB) compared to 0.85 for explicit models,[25] suggesting that implicit solvent models should be used with caution. In two studies, the reference interaction site models (RISM[26–28] and 3D-RISM[29–31]) have been used for HFE calculations of drug-like molecules.[22,32] However, even though statistical mechanical theories developed for molecular liquids by using the distribution functions of the system are compute-efficient compared to MD, their HFE results are poorer due to the approximations in theories.[2,33,34]

In order to compute the accurate hydration free energies of small molecules via MD, many researchers have focused

a)Author to whom correspondence should be addressed: david.vanderspoel@icm.uu.se.

on equilibrium methods such as BAR[35] and MBAR.[36] These methods are based on data collected from the equilibrium simulations as well as free-energy perturbations. On the other hand, non-equilibrium methods such as the one-sided Jarzynski Equality (JE)[37] and two-sided Crook Gaussian Intersection (CGI)[38–41] methods have been proposed to be compute-efficient alternatives for estimating the hydration free energies of small molecules. Here, we calculated the HFE of 34 small neutral organic molecules with the experimental range from +2 kcal to −16 kcal/mol with both the two-sided CGI, BAR, and MBAR methods and the one-sided JE using MD. The calculated HFE results are compared with MBAR results and experimental values from the literature.

II. METHODS

The free energy corresponds to the amount of work that a system can perform. Free energy differences between two equilibrium states of a system can be calculated, using MD simulations, with the Jarzynski equality (JE), the Crooks Gaussian Intersection (CGI) relation, or the Bennett Acceptance Ratio (BAR) method. These methods are explained briefly below:

A. Jarzynski Equality

JE computes the free energy difference ΔG_{AB} between two equilibrium states from the exponential average of the work W performed in non-equilibrium transitions from one state to the other. It requires that the transition be started from an equilibrium state. The energy ΔG_{AB} is determined from

$$e^{-\beta \Delta G_{AB}} = \left\langle e^{-\beta W} \right\rangle. \tag{1}$$

The work done on the system along the variable λ which regulates the potential energy of the system can be obtained by

$$W = \int_0^1 \frac{\partial H}{\partial \lambda} d\lambda. \tag{2}$$

In Eq. (2), instantaneous $\partial H/\partial \lambda$ values are integrated. The resulting work includes contributions from the free energy difference between the states and the dissipated work along the transition path. However, it has been reported that the Jarzynski estimator is biased by the finite number of the trajectories.[42] For a small system in a near-equilibrium state, the magnitude of this bias $B_J(N)$ can be estimated with an empirical relationship proposed by Gore *et al.*[42]

$$B_J(N) = \frac{\bar{W}_{dis}}{N^{\alpha}}, \tag{3}$$

where N is the number of trajectories ($N = 50$ in our case), $\bar{W}$ is the mean dissipated work which is defined as the difference between the work and the equilibrium free energy difference, α is a decreasing function of $\bar{W}$, and

$$\bar{W}_{dis} = \langle W \rangle - \Delta G = \frac{1}{2} \beta \sigma_w^2, \tag{4}$$

where $\beta = (RT)^{-1}$, T is the temperature and R is the gas constant ($R = 0.008\,31$ kJ/mol K), and σ_w^2 is the variance of the work distribution. Gore *et al.* have reported that bias needs a function α, which depends on a parameter C. The α function is expressed as

$$\alpha = \frac{ln\left[2\beta C \bar{W}_{dis}\right]}{ln\left[C(e^{2\beta \bar{W}_{dis}} - 1)\right]}. \tag{5}$$

The bias-corrected JE is defined as

$$\Delta G_J = \Delta G - B_J(N). \tag{6}$$

The variance for bias-corrected JE is given by

$$\sigma_J^2(N) = \frac{\sigma_w^2}{N^{\alpha v}}. \tag{7}$$

For bias calculation, Eq. (5) was used with $C = 15$, but for the variance, $\alpha v = \alpha(C = 50)$ based on the work by Gore *et al.*[42] Finally, the mean square error (MSE) for the bias-corrected Jarzynski estimator is defined as

$$MSE = B_J^2(N) + \sigma_J^2(N). \tag{8}$$

Here, we used three variations of the bias-corrected JE: Jarzynski Forward (JF), in which the simulation is started in the hydrated state, Jarzynski Backward (JB), starting from the gas state (decoupled), and Jarzynski Mean (JM), taking the mean of JF and JB.

B. Crooks Gaussian Intersection

The Crooks Fluctuation Theorem (CFT) is an equation based on the ratio of the work distributions of two-sided transitions (that are started from equilibrium ensembles), moving from state A to state B and from state B to state A, with dissipated work involved in the transformation. The free energy difference can be computed from the CFT as the work value W where both distributions overlap $P_f(W) = P_b(-W)$,

$$\frac{P_f(W)}{P_b(-W)} = e^{\beta(W - \Delta G)}. \tag{9}$$

Plotting the logarithm of the left side of Eq. (9) against the work values obtained from the non-equilibrium transitions gives a line with a slope β. This line intercepts the work axis at a value equal to ΔG.

Goette and Grubmüller derived a CGI estimator which is based on Gaussian approximation and showed that CGI yields accurate free energy estimations.[41] The CGI method is mathematically expressed as

$$\Delta G = \frac{\frac{\langle W_f \rangle_{n_f}}{\sigma_f^2} - \frac{-\langle W_b \rangle_{n_b}}{\sigma_b^2} \mp \sqrt{\frac{1}{\sigma_f^2 \sigma_b^2}\left(\left\langle W_f \right\rangle_{n_f} + \langle W_b \rangle_{n_b}\right)^2 + 2\left(\frac{1}{\sigma_f^2} - \frac{1}{\sigma_b^2}\right) ln \frac{\sigma_b}{\sigma_f}}}{\frac{1}{\sigma_f^2} - \frac{1}{\sigma_b^2}}, \tag{10}$$

where $\langle W_f \rangle_{n_f}$ and $\langle W_b \rangle_{n_b}$ are the work averages in the forward and backward directions, respectively. n_f and n_b are the number of the transitions of the respective directions, σ_f and σ_b are the standard deviations of the respective Gaussian functions, and σ_f^2 and σ_b^2 are the variances of the work distributions for the transitions in both directions.

For the calculations of HFE with CGI, we used the work distributions of two-sided (forward and backward) transitions based on the Gaussian approximation. The HFE ΔG is where the two distributions intercept (Fig. 1).

C. Bennett Acceptance Ratio

The BAR method estimates the free energy difference between two states A and B based on the data obtained from the simulations of both states simultaneously. The free energy difference between two states is given as

$$\Delta G_{AB} = \frac{1}{\beta} ln \frac{\langle f(H_A(p,q) - H_B(p,q)) + C \rangle_B}{\langle f(H_B(p,q) - H_A(p,q)) - C \rangle_A} + C, \quad (11)$$

where $\beta = 1/k_B T$, k_B is the Boltzmann constant, and T is the temperature. The subscripts A and B in Eq. (11) denote that the ensemble averages < > are calculated from the trajectories of the initial (A) and final (B) states, respectively. The symbols f and H represent the Fermi function $f(x) = \frac{1}{1+\exp(\beta x)}$ and the Hamiltonian of a system that consists of a kinetic and a potential energy component $H(p,q) = K(p) + U(q)$, respectively. The unknown constant C is found numerically from

$$C = \frac{1}{\beta} \left(\frac{Q_A}{Q_B} \frac{n_A}{n_B} \right). \quad (12)$$

Here n_A and n_B are the number of samples in each state and Q denotes the corresponding partition function. The value of C obtained from Eq. (12) yields the free energy difference ΔG_{AB},

$$\Delta G_{AB} = -\frac{1}{\beta} ln \frac{n_B}{n_A} + C. \quad (13)$$

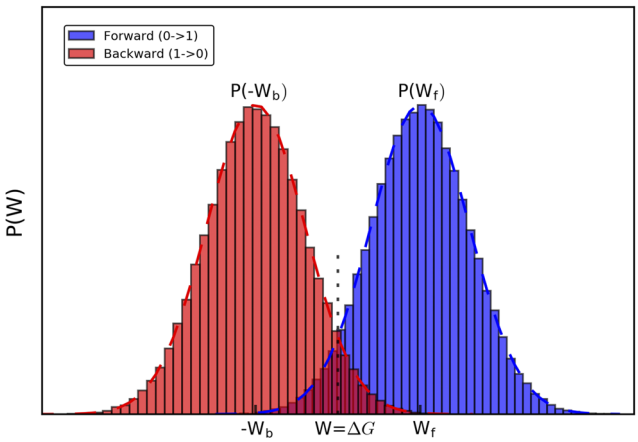

FIG. 1. Schematic representation of the non-equilibrium HFE calculation for switching from the equilibrium state A to another state B for a forward (A → B) and a backward (B → A) process with CGI. W_f and W_b are the means of the work distributions of the forward [$P(W_f)$] and backward state [$P(-W_b)$], respectively. ΔG is the intercept of the two distributions.

To assess data from multiple states, Shirts and Chodera[36] extended the method to the Multistate Bennett Acceptance Ratio (MBAR), using a maximum likelihood formulation.

III. SIMULATION DETAILS

All simulations were run using the MD software package Groningen Machine for Chemical Simulation (GROMACS) (version 2016.2).[43–47] The force field models and the initial coordinates of compounds were taken from the FreeSolv database.[7,8] In all simulations, we used a leap-frog stochastic dynamics integrator[48] with the AMBER99SB-ILDN force field,[49] TIP3P water model,[50] the Linear Constraint Solver (LINCS) algorithm[51] for hydrogen bond constraints, and SETTLE algorithm[52] to keep the water bonds and angle rigid. The equations of motion were integrated using a time step of 2 fs. Temperature coupling was performed using Langevin dynamics[48,53] with a coupling constant of 1.0 ps and a reference temperature of 298.15 K. To establish and maintain a pressure of 1 bar, the Berendsen barostat[54] was used during the equilibrium stage of the simulations and the Parrinello-Rahman barostat[55] was used during production simulations, with a time constant of 1.0 ps and a compressibility of 4.5×10^{-5} bar^{-1}. The particle mesh Ewald (PME) algorithm[56] was used for electrostatic interactions with a switching distance of 1.2 nm, a grid spacing of 0.12 nm, and an interpolation order of 6 for long range electrostatics. For the van der Waals interactions, a cutoff of 1.1 nm was used. A softcore potential[57] was used during free energy calculations with parameters α = 0.3, σ = 0.25, and p = 1.

For the HFE calculations, the non-equilibrium transition simulations were conducted based on a non-equilibrium fast-growth thermodynamic integration (FGTI) protocol. First, 10 ns simulations were performed in the equilibrium states A (coupled) and B (decoupled). Then, from the first 5 ns of the simulations, 50 snapshots were extracted and used to run short non-equilibrium simulations of 100 ps each, in which the coupling with the environment was inverted by switching lambda from 0 to 1 (forward: decoupling, $A \rightarrow B$) and from 1 to 0 (backward: coupling, $B \rightarrow A$) taking increments/decrements of $\pm 2 \times 10^{-5}$. For the analysis of the forward and backward simulations, the pmx tool was used.[58] The uncertainties associated with each method were obtained with 1000 bootstrap iterations except for MBAR. Since MBAR presents an underestimate of the true uncertainty in the values due to finite sampling,[59] we performed three replicates of the simulations to obtain a more realistic uncertainty.

IV. RESULTS AND DISCUSSION

HFE calculations were performed on the 34 small neutral organic molecules given in Table I with the experimental range of free energies from +2 kcal to −16 kcal/mol using six different methods, JF, JB, JM, CGI, BAR, and MBAR, based on MD simulations. The MBAR results presented here are virtually identical to the FreeSolv values [Pearson R^2 = 99.95%, root mean square deviation (RMSD) 0.02 kcal/mol]. Here, the main focus is on comparing the statistical efficiency of the methods,

TABLE I. The hydration free energies from the first 5 ns of the trajectories of all the compounds studied. Energies are in kcal/mol. $\Delta G_{Exp.}$ values are taken from the FreeSolv database.[7,8] Statistics are computed with respect to experimental data, mean absolute error (MAE), root mean square error (RMSE), and Pearson correlation coefficient R.

IUPAC name	ΔG_{JF}	ΔG_{JB}	ΔG_{JM}	ΔG_{CGI}	ΔG_{BAR}	ΔG_{MBAR}	$\Delta G_{Exp.}$
Octane	2.97 ± 0.18	2.89 ± 0.22	2.93 ± 0.15	3.05 ± 0.10	3.00 ± 0.14	3.08 ± 0.04	2.88
Heptane	2.57 ± 0.14	3.37 ± 0.13	2.97 ± 0.10	2.94 ± 0.09	2.95 ± 0.10	2.94 ± 0.02	2.67
Hexane	2.87 ± 0.22	2.88 ± 0.15	2.88 ± 0.14	2.94 ± 0.10	2.83 ± 0.10	2.82 ± 0.01	2.48
n-pentane	2.69 ± 0.18	2.66 ± 0.16	2.68 ± 0.12	2.72 ± 0.09	2.67 ± 0.08	2.69 ± 0.03	2.30
n-butane	2.29 ± 0.08	2.41 ± 0.22	2.35 ± 0.11	2.65 ± 0.07	2.58 ± 0.07	2.57 ± 0.02	2.10
Benzene	−1.01 ± 0.12	−0.71 ± 0.10	−0.86 ± 0.08	−0.78 ± 0.09	−0.86 ± 0.08	−0.80 ± 0.03	−0.90
Toluene	−0.69 ± 0.15	−0.71 ± 0.12	−0.70 ± 0.10	−0.76 ± 0.13	−0.74 ± 0.08	−0.76 ± 0.02	−0.90
Chloroform	0.27 ± 0.11	0.28 ± 0.11	0.28 ± 0.07	0.34 ± 0.08	0.28 ± 0.07	0.26 ± 0.01	−1.08
Styrene	−1.26 ± 0.11	−1.01 ± 0.14	−1.14 ± 0.09	−1.03 ± 0.08	−1.08 ± 0.09	−1.07 ± 0.01	−1.24
Octanal	−2.74 ± 0.13	−2.74 ± 0.18	−2.74 ± 0.11	−2.50 ± 0.13	−2.62 ± 0.11	−2.56 ± 0.01	−2.29
Pentyl acetate	−2.34 ± 0.40	−2.55 ± 0.14	−2.45 ± 0.22	−2.81 ± 0.11	−2.84 ± 0.13	−2.79 ± 0.07	−2.51
Tetrahydrofuran	−2.12 ± 0.26	−2.04 ± 0.09	−2.08 ± 0.13	−2.17 ± 0.07	−2.22 ± 0.06	−2.16 ± 0.02	−3.47
Acetone	−3.55 ± 0.08	−3.44 ± 0.11	−3.50 ± 0.07	−3.44 ± 0.07	−3.47 ± 0.07	−3.55 ± 0.02	−3.80
Acetonitrile	−2.73 ± 0.09	−2.83 ± 0.07	−2.78 ± 0.06	−2.88 ± 0.08	−2.81 ± 0.05	−2.79 ± 0.02	−3.88
1-Octanol	−2.65 ± 0.24	−2.95 ± 0.19	−2.80 ± 0.15	−2.78 ± 0.13	−2.86 ± 0.14	−2.71 ± 0.01	−4.09
Ethanol	−3.50 ± 0.07	−3.43 ± 0.12	−3.47 ± 0.07	−3.38 ± 0.07	−3.41 ± 0.05	−3.42 ± 0.03	−5.00
1,4-Dioxane	−4.30 ± 0.13	−4.12 ± 0.14	−4.21 ± 0.10	−4.13 ± 0.09	−4.18 ± 0.07	−4.21 ± 0.05	−5.06
Methanol	−3.54 ± 0.06	−3.51 ± 0.06	−3.53 ± 0.04	−3.44 ± 0.06	−3.50 ± 0.04	−3.48 ± 0.02	−5.10
p-cresol	−5.84 ± 0.12	−5.65 ± 0.16	−5.75 ± 0.09	−5.56 ± 0.09	−5.67 ± 0.09	−5.55 ± 0.01	−6.13
Benzoquinone	−7.01 ± 0.12	−6.72 ± 0.14	−6.87 ± 0.09	−6.86 ± 0.09	−6.83 ± 0.08	−6.95 ± 0.03	−6.50
Morpholine	−6.30 ± 0.14	−5.96 ± 0.11	−6.13 ± 0.09	−6.02 ± 0.09	−6.13 ± 0.07	−6.13 ± 0.02	−7.17
1-Methylpiperazine	−8.49 ± 0.19	−8.34 ± 0.20	−8.42 ± 0.13	−8.26 ± 0.10	−8.34 ± 0.12	−8.19 ± 0.03	−7.77
2-naphthol	−7.95 ± 0.21	−8.37 ± 0.33	−8.16 ± 0.19	−7.68 ± 0.10	−7.82 ± 0.11	−7.83 ± 0.03	−8.11
Sulfolane	−10.09 ± 0.14	−9.63 ± 0.11	−9.86 ± 0.09	−9.83 ± 0.09	−9.86 ± 0.09	−9.58 ± 0.04	−8.61
Captan	−8.45 ± 0.43	−8.24 ± 0.19	−8.35 ± 0.23	−8.68 ± 0.16	−8.67 ± 0.18	−8.62 ± 0.03	−9.01
Methyl paraben	−9.61 ± 0.20	−9.99 ± 0.17	−9.80 ± 0.12	−9.76 ± 0.11	−9.81 ± 0.14	−9.81 ± 0.04	−9.51
N-methylacetamide	−8.32 ± 0.16	−8.31 ± 0.11	−8.32 ± 0.10	−8.45 ± 0.07	−8.39 ± 0.06	−8.30 ± 0.03	−10.00
Simazine	−10.11 ± 0.60	−11.12 ± 0.21	−10.62 ± 0.33	−10.87 ± 0.14	−10.89 ± 0.15	−10.70 ± 0.07	−10.22
Benzamide	−10.57 ± 0.10	−10.36 ± 0.16	−10.47 ± 0.09	−10.23 ± 0.09	−10.34 ± 0.09	−10.44 ± 0.03	−11.00
Terbacil	−13.33 ± 0.23	−13.63 ± 0.18	−13.48 ± 0.15	−13.60 ± 0.14	−13.62 ± 0.15	−13.73 ± 0.03	−11.14
Glycerol	−10.31 ± 0.40	−10.73 ± 0.13	−10.52 ± 0.21	−10.82 ± 0.10	−10.84 ± 0.10	−10.76 ± 0.08	−13.43
5-Trifluoromethyluracil	−16.83 ± 0.34	−17.15 ± 0.13	−16.99 ± 0.18	−17.31 ± 0.10	−17.30 ± 0.11	−17.33 ± 0.06	−15.46
6-Chlorouracil	−15.15 ± 0.24	−15.23 ± 0.15	−15.19 ± 0.13	−15.25 ± 0.10	−15.24 ± 0.10	−15.16 ± 0.05	−15.83
5-fluorouracil	−16.28 ± 0.26	−16.38 ± 0.09	−16.33 ± 0.14	−16.55 ± 0.10	−16.57 ± 0.09	−16.34 ± 0.02	−16.92
Simulation time (ns)	10	10	20	20	20	100	
Average RMSD	0.20	0.15	0.13	0.10	0.10	0.03	
Statistical Efficiency $\mathcal{E}$	2.6	4.6	3.2	5.3	5.4	11.6	
Statistics							
MAE	0.78 ± 0.12	0.84 ± 0.11	0.79 ± 0.12	0.83 ± 0.11	0.80 ± 0.11	0.81 ± 0.12	
RMSE	1.05 ± 0.32	1.07 ± 0.28	1.05 ± 0.30	1.05 ± 0.28	1.04 ± 0.28	1.06 ± 0.30	
Pearson R	0.98 ± 0.01	0.98 ± 0.01	0.98 ± 0.01	0.98 ± 0.01	0.98 ± 0.01	0.98 ± 0.01	

but first, we establish the correctness and convergence of the calculations.

To assess the convergence of the initial 10 ns trajectories, both in the coupled and the decoupled state, each was divided into two parts of 5 ns. The average root mean square deviation (RMSD) for both parts was calculated by dividing the first and last 5 ns of the MD trajectory into five blocks of 1 ns, computing the mean values in each block, and then calculating the standard deviation of the mean values from the five blocks. The similarity between the average RMSDs of the first and second parts of the trajectories (Fig. S1) suggests convergence in both parts of the simulations in the coupled state, where van der Waals and Coulomb interactions between the small molecule and the surrounding water molecules are on. This then suggests that only 5 ns of simulation should be sufficient for the selection of starting points for the non-equilibrium HFE calculations. To check the effect of simulation time on the accuracy of HFE, we calculated the hydration free energies of all the compounds studied for both trajectory parts as well. For the last 5 ns of the trajectories, the calculated hydration free energies are given in Table S1, but the results and figures presented in the main text were obtained using the first 5 ns of the equilibrated trajectories of 10 ns. The difference between the results obtained from the first and second half of

the trajectory is small, as can be seen comparing Tables I and S1. The results of HFE without bias correction, bias, variance, and mean square error (MSE) for JF and JB are presented in Tables S2 and S3.

Detailed results of the HFE estimates of the first 5 ns and the last 5 ns of the trajectories obtained using the CGI method are provided as the supplementary material (Figs. S2–S35 and Figs. S36–S69, respectively). It is noted that in those supplementary graphs, the sign of ΔG should be inverted since HFE is ΔG (B → A) = G (state A) − G (state B), while CGI is a specific estimator for the ΔG (A → B) = G (state B) – G (state A) based on the CFT. As can be seen from the joint set of CGI plots (Figs. S2–S69), for all the molecules, the work distributions of the forward and backward states exhibit a large overlap, suggesting that the equilibrated A and B states have converged well.

Table I lists the computational predictions of JF, JB, JM, CGI, BAR, and MBAR together with the experimental data for all the molecules studied. From this table, it is possible to assess how well the computational predictions correspond to the experimental values and which methods give similar or different results. Table I shows that all methods produce the same mean absolute error (MAE) and root mean square error (RMSE) from experiment, which is a prerequisite for further analysis. It is apparent that the error bars in the individual HFE are larger for JF, JB, JM, CGI, and BAR than for MBAR. The error bars are related to the amount of simulation time per molecule of the employed methods (Table I). The MBAR results involve a total of 20 windows (intermediate λ states) of 5 ns for each molecule, thus costing 100 ns for each molecule. In this study, we used equilibrium trajectories of 5 ns in both the coupled and the decoupled states. From these trajectories of the simulations, 50 snapshots were extracted and non-equilibrium transitions (simulations) were performed for 100 ps each. The cost of both the equilibrium and non-equilibrium simulations is 5 ns, with a total of 10 ns for each molecule using either of the one-sided methods, and double that for the two-sided methods (Table I).

To assess the accuracy of the calculated HFE for each method and see whether the size of the deviations depends on the size of the HFE and/or the nature of the compound, the deviations from the experimental values are shown in Fig. 2, with gray bands denoting the range of absolute errors within 1 and 2 kcal/mol. The figure shows that ΔΔG free energy differences of all the methods with experimental values are within 1 kcal/mol for 22 molecules, within 2 kcal/mol for 10 molecules, and within 3 kcal/mol for 2 molecules. From Fig. 2, it appears that both one- and two-sided methods employed in this study can be used to predict the experimental HFE values, but the use of one-sided methods, in particular, JF, leads to slightly larger systematic errors. Although the methods yield different RMSE, the deviation from the "true" value defined by the force field can be assumed to be stochastic with an expectation value of zero at least for the two-sided methods. As a result, the deviation from experiment is identical for these methods. The correction for the sampling bias[42] compensates this to a

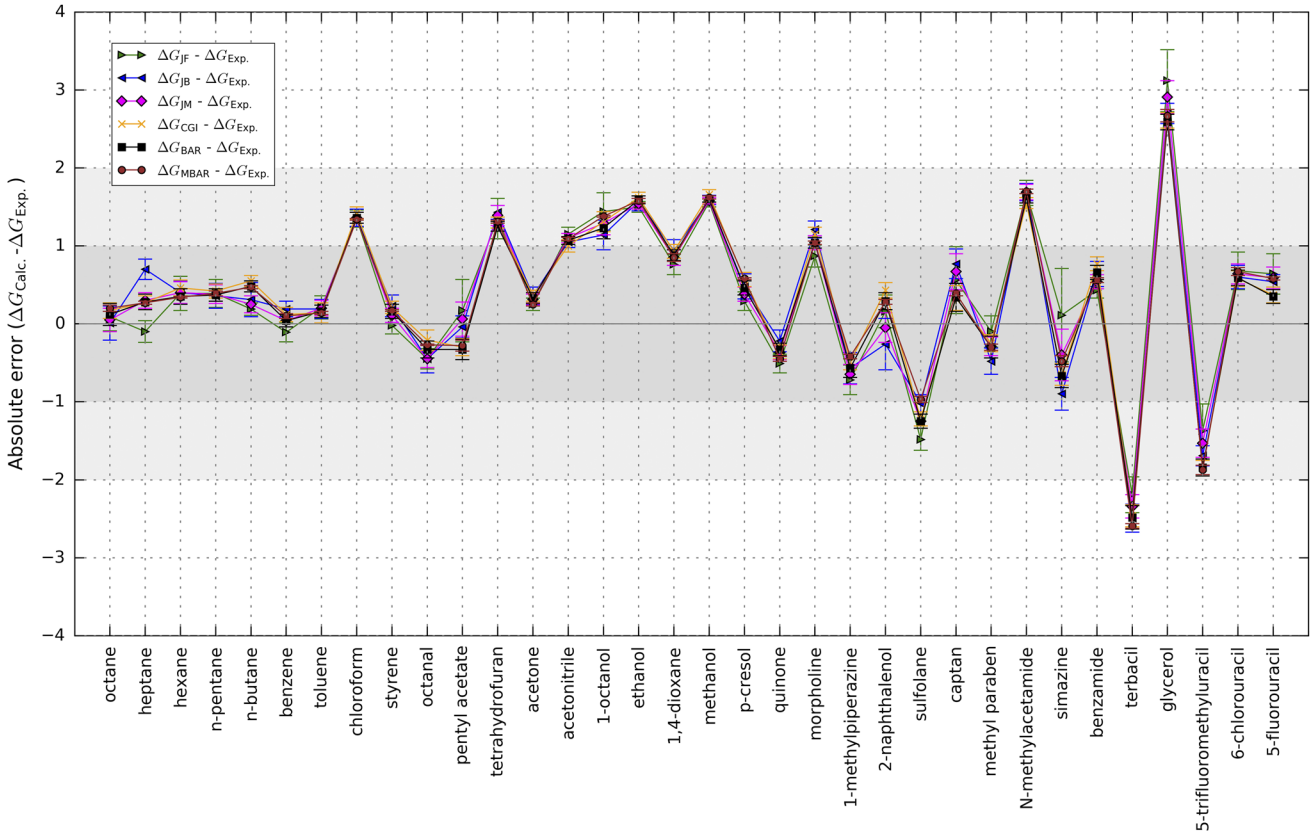

FIG. 2. Absolute errors between calculated and experimental hydration free energies. The darker and lighter gray shaded areas mark a deviation of 1 kcal/mol and 2 kcal/mol, respectively. Lines are drawn between neighboring points to guide the eye and stress the profiles in the deviations for comparing the different methods.

large extent. This means there is no established way to correct the JF or JB methods to obtain the true HFE ("true" as in the value inherent to the force field used). However, averaging JF and JB into JM results in a reliable estimate of the HFE, as proposed by Collin *et al.*[60]

To investigate force-field consistency, a linear regression was performed of the computed values against the MBAR values for each of the methods in Fig. 3. The correlation coefficients between the computed hydration free energies with JF, JB, JM, CGI, BAR, and MBAR have R of 1.0.

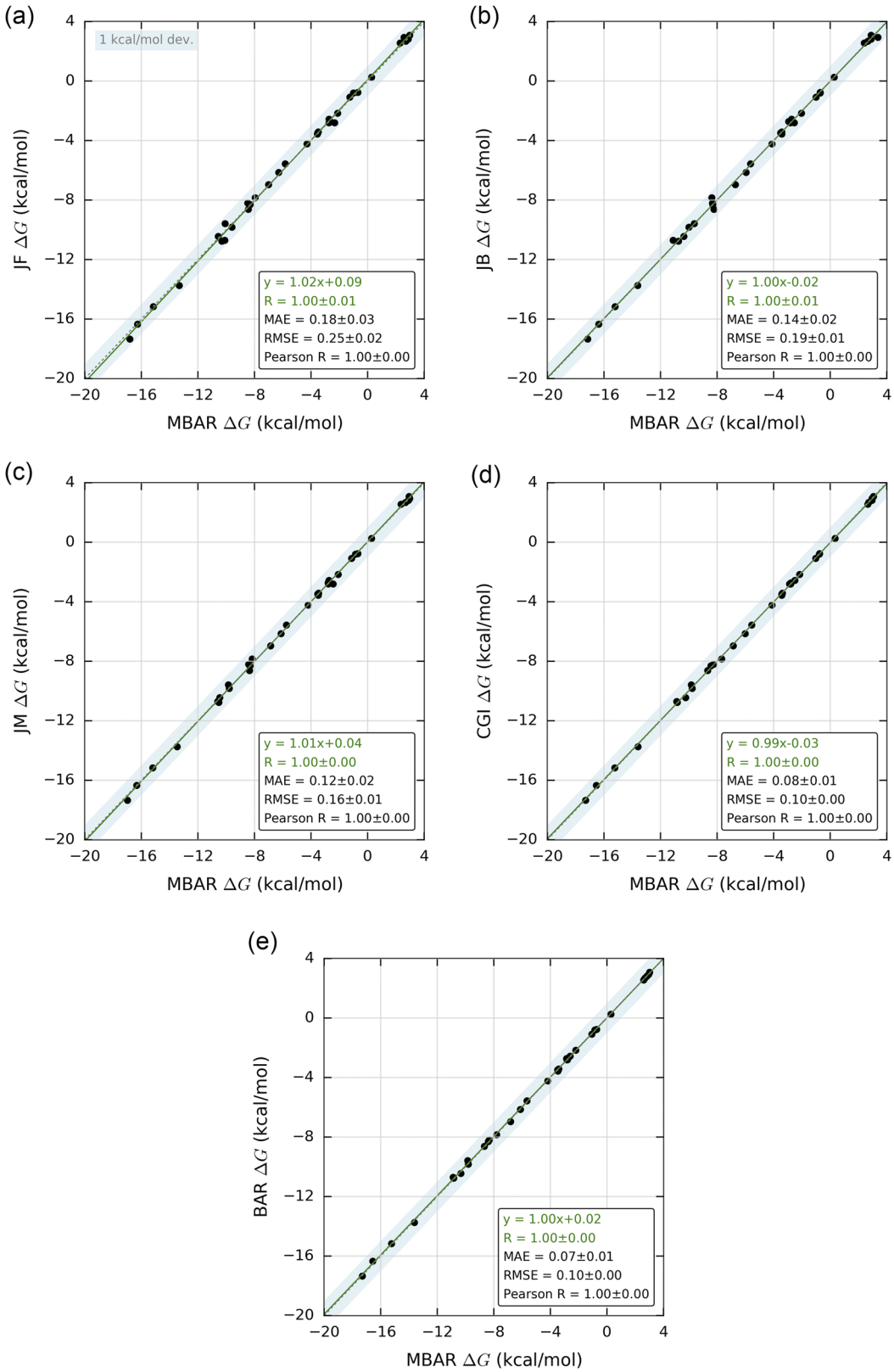

FIG. 3. Scatter and correlation plots of calculated versus MBAR free energies. Hydration free energies from (a) JF, (b) JB, (c) JM, (d) CGI, and (e) BAR.

Additionally, the dataset of each method was compared statistically with MBAR values. The absolute error (MAE) and the root mean square error (RMSE) for all methods are given in Fig. 3. These show that JM, CGI, and BAR yield results that are more consistent with MBAR than the one-sided JF and JB.

Finally, we return to Table I and evaluate the statistical efficiency ε of each of the methods, defined as[61]

$$\varepsilon = \frac{1}{T\langle MSD\rangle},$$

where T is the simulation time that is multiplied by the average variance (MSD) in the HFE calculations. Here, the simulation time is used as a proxy for the amount of information in the simulation. This yields a number ε for each method that can be used as follows: for a desired ⟨RMSD⟩, the simulation time T needs to be $1/(\varepsilon\langle RMSD\rangle^2)$. This makes sense if we assume independent statistical sampling since in this case, the RMSD decreases as the square root of the sampling time. Table S1 shows that the ε are reproducible since they are roughly the same as in Table I. The MBAR method has more than double the efficiency of CGI and BAR, both of which have slightly higher efficiency than either JM, JF, or JB. In order for JM to obtain the same RMSD as MBAR, which is considered the best and most reliable estimator,[7,62] approximately 13 times more sampling would be needed, annulling any perceived computational efficiency gains of the non-equilibrium method. It should be noted that the absolute values of ε are dependent on the dataset being considered and that the amount of information may differ somewhat between different methods. There is no *a priori* reason to believe, however, that the relative efficiencies would be different for other datasets. It should be noted that the accuracy of non-equilibrium free energy methods depends on a number of settings, such as the size of molecule, the speed of (de)coupling, and the simulation time of the equilibrium trajectories and the non-equilibrium transitions. It is likely that the larger/flexible molecule, faster (de)coupling, and shorter simulation time would lead to higher variance due to the large and/or fast perturbations of the molecule, especially in the decoupled state, and hence lower statistical efficiency as what is found here. Paliwal and Shirts performed an extensive statistical analysis of free energy methods, including thermodynamic integration, BAR and MBAR. In line with our results, they concluded that MBAR is the most efficient method in terms of the lowest attainable error;[62] however, they did not take into account the amount of sampling needed to obtain this efficiency. A point of concern in our analysis is whether the errors in the equation for ε are accurate or indeed comparable between different methods. The error estimators used here are considered to be state of the art for all methods for BAR and MBAR as well as for the non-equilibrium methods.[62] Should better estimators be derived, the analysis may need to be redone.

V. CONCLUSION

Hydration free energies of 34 small neutral organic molecules were computed using the non-equilibrium free energy calculation methods JF, JB, JM, BAR, and CGI and the results compared with the equilibrium method MBAR and experimental values. Comparison of the different methods shows that all the non-equilibrium one- and two-sided methods reproduce the HFE results from the equilibrium two-sided method MBAR reasonably well. A careful comparison of the efficiency ε of the methods as expressed in average RMSD divided by simulation time yields that ε(MBAR) > ε(BAR) ≈ ε(CGI) > ε(JX), where JX is any of the Jarzynski methods.

It should be noted that all the molecules studied are small. Large and/or flexible molecules may suffer from convergence issues due to the large perturbations caused by their internal motions and local roto-translational motions in the decoupled state, even with two-sided methods. Since free energy is a state function, therefore, splitting a large perturbation into series of smaller ones by multistage free energy perturbation calculations using multiple λ values is recommended.[63]

SUPPLEMENTARY MATERIAL

See supplementary material for figures of the predicted hydration free energies of all the compounds with the CGI method.

ACKNOWLEDGMENTS

The Swedish research council is acknowledged for financial support to D.v.d.S. (Grant No. 2013-5947) and for a grant of computer time (No. SNIC2017-12-41) through the High Performance Computing Center North in Umeå, Sweden, and the PDC Center for High Performance Computing at the Royal Institute of Technology, Stockholm, Sweden.

[1] G. Schirò, Y. Fichou, F. X. Gallat, K. Wood, F. Gabel, M. Moulin, M. Härtlein, M. Heyden, J. P. Colletier, A. Orecchini, A. Paciaroni, J. Wuttke, D. J. Tobias, and M. Weik, Nat. Commun. **6**, 6490 (2015).
[2] T. Yoshidome, T. Ekimoto, N. Matubayasi, Y. Harano, M. Kinoshita, and M. Ikeguchi, J. Chem. Phys. **142**, 175101 (2015).
[3] N. A. Mohamed, R. T. Bradshaw, and J. W. Essex, J. Comput. Chem. **37**, 2749 (2016).
[4] D. L. Mobley, K. A. Dill, and J. D. Chodera, J. Phys. Chem. B **112**, 938 (2008).
[5] D. L. Mobley, C. I. Bayly, M. D. Cooper, M. R. Shirts, and K. A. Dill, J. Chem. Theory Comput. **5**, 350 (2009).
[6] C. H. Bennett, J. Comput. Phys. **22**, 245 (1976).
[7] G. D. R. Matos, D. Y. Kyu, H. H. Loeffler, J. D. Chodera, M. R. Shirts, and D. L. Mobley, J. Chem. Eng. Data **62**, 1559 (2017).
[8] D. L. Mobley and J. P. Guthrie, J. Comput.-Aided Mol. Des. **28**, 711 (2014).
[9] M. R. Shirts and V. S. Pande, J. Chem. Phys. **122**, 134508 (2005).
[10] B. Hess and N. F. A. van der Vegt, J. Phys. Chem. B **110**, 17616 (2006).
[11] Y. Deng and B. Roux, J. Phys. Chem. B **108**, 16567 (2004).
[12] A. Villa and A. E. Mark, J. Comput. Chem. **23**, 548 (2002).
[13] H. Zhang, Y. Jiang, H. Yan, C. Yin, T. Tan, and D. Van Der Spoel, J. Phys. Chem. Lett. **8**, 2705 (2017).
[14] A. Nicholls, D. L. Mobley, J. P. Guthrie, J. D. Chodera, C. I. Bayly, M. D. Cooper, and V. S. Pande, J. Med. Chem. **51**, 769 (2008).
[15] D. L. Mobley, E. Dumont, J. D. Chodera, and K. A. Dill, J. Phys. Chem. B **111**, 2242 (2007).
[16] G. J. Rocklin, D. L. Mobley, K. A. Dill, and P. H. Hünenberger, J. Chem. Phys. **139**, 184103 (2013).
[17] J. W. Kaus, L. T. Pierce, R. C. Walker, and J. A. McCammon, J. Chem. Theory Comput. **9**, 4131 (2013).
[18] P. Mikulskis, S. Genheden, and U. Ryde, J. Chem. Inf. Model. **54**, 2794 (2014).
[19] W. Clark Still, A. Tempczyk, R. C. Hawley, and T. Hendrickson, J. Am. Chem. Soc. **112**, 6127 (1990).

[20]B. Honig and A. Nicholls, Science **268**, 1144 (1995).
[21]N. A. Baker, D. Bashford, and D. A. Case, *New Algorithms Macromolecular Simulation* (Springer-Verlag, Berlin/Heidelberg, 2006), pp. 263–295.
[22]J. Johnson, D. A. Case, T. Yamazaki, S. Gusarov, A. Kovalenko, and T. Luchko, J. Phys.: Condens. Matter **28**, 344002 (2016).
[23]H. Zhang, T. Tan, and D. van der Spoel, J. Chem. Theory Comput. **11**, 5103 (2015).
[24]H. Zhang, C. Yin, H. Yan, and D. Van Der Spoel, J. Chem. Inf. Model. **56**, 2080 (2016).
[25]J. Zhang, H. Zhang, T. Wu, Q. Wang, and D. Van Der Spoel, J. Chem. Theory Comput. **13**, 1034 (2017).
[26]H. C. Andersen and D. Chandler, J. Chem. Phys. **57**, 1918 (1972).
[27]F. Hirata, P. J. Rossky, and B. Montgomery Pettitt, J. Chem. Phys. **78**, 4133 (1983).
[28]J. Perkyns and B. M. Pettitt, J. Chem. Phys. **97**, 7656 (1992).
[29]A. Kovalenko and F. Hirata, J. Chem. Phys. **110**, 10095 (1999).
[30]D. Beglov and B. Roux, J. Phys. Chem. B **101**, 7821 (1997).
[31]E. L. Ratkova, D. S. Palmer, and M. V. Fedorov, Chem. Rev. **115**, 6312 (2015).
[32]D. S. Palmer, V. P. Sergiievskyi, F. Jensen, and M. V. Fedorov, J. Chem. Phys. **133**, 044104 (2010).
[33]Y. Karino, M. V. Fedorov, and N. Matubayasi, Chem. Phys. Lett. **496**, 351 (2010).
[34]M. Kinoshita, J. Chem. Phys. **128**, 024507 (2008).
[35]M. R. Shirts, E. Bair, G. Hooker, and V. S. Pande, Phys. Rev. Lett. **91**, 140601 (2003).
[36]M. R. Shirts and J. D. Chodera, J. Chem. Phys. **129**, 124105 (2008).
[37]C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[38]G. E. Crooks, J. Stat. Phys. **90**, 1481 (1998).
[39]G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).
[40]R. Chelli, S. Marsili, A. Barducci, and P. Procacci, J. Chem. Phys. **126**, 044502 (2007).
[41]M. Goette and H. Grubmüller, J. Comput. Chem. **30**, 447 (2009).
[42]J. Gore, F. Ritort, and C. Bustamante, Proc. Natl. Acad. Sci. U. S. A. **100**, 12564 (2003).
[43]H. J. C. Berendsen, D. van der Spoel, and R. van Drunen, Comput. Phys. Commun. **91**, 43 (1995).
[44]E. Lindahl, B. Hess, and D. van der Spoel, J. Mol. Model. **7**, 306 (2001).
[45]D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, J. Comput. Chem. **26**, 1701 (2005).
[46]B. Hess, C. Kutzner, D. Van Der Spoel, and E. Lindahl, J. Chem. Theory Comput. **4**, 435 (2008).
[47]S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. Van Der Spoel, B. Hess, and E. Lindahl, Bioinformatics **29**, 845 (2013).
[48]W. F. van Gunsteren and H. J. C. Berendsend, Mol. Simul. **1**, 173 (1988).
[49]K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, Proteins: Struct., Funct., Bioinf. **78**, 1950 (2010).
[50]W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[51]B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, J. Comput. Chem. **18**, 1463 (1997).
[52]S. Miyamoto and P. A. Kollman, J. Comput. Chem. **13**, 952 (1992).
[53]N. Goga, A. J. Rzepiela, A. H. De Vries, S. J. Marrink, and H. J. C. Berendsen, J. Chem. Theory Comput. **8**, 3637 (2012).
[54]H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. Dinola, and J. R. Haak, J. Chem. Phys. **81**, 3684 (1984).
[55]M. Parrinello and A. Rahman, J. Appl. Phys. **52**, 7182 (1981).
[56]U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, J. Chem. Phys. **103**, 8577 (1995).
[57]T. C. Beutler, A. E. Mark, R. C. van Schaik, P. R. Gerber, and W. F. van Gunsteren, Chem. Phys. Lett. **222**, 529 (1994).
[58]V. Gapsys, S. Michielssens, D. Seeliger, and B. L. De Groot, J. Comput. Chem. **36**, 348 (2015).
[59]M. Fajer, R. Swift, and J. McCammon, J. Comput. Chem. **30**, 1719 (2009).
[60]D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco, and C. Bustamante, Nature **437**, 231 (2005).
[61]R. A. Fisher, Philos. Trans. R. Soc., A **222**, 309 (1922).
[62]H. Paliwal and M. R. Shirts, J. Chem. Theory Comput. **7**, 4115 (2011).
[63]C. Chipot and A. Pohorille, *Free Energy Calculations* (Springer-Verlag, Berlin/Heidelberg, 2007).